(12) UK Patent Application (19) GB (11) 2 235 815 (13) A

(43) Date of A publication 13.03.1991

(21) Application No 9018790.7

(22) Date of filing 28.08.1990

(30) Priority data
(31) 402133 (32) 01.09.1989 (33) US
469455 20.03.1990

(71) Applicant
Compact Video Group Inc

(Incorporated in the USA - California)

2813 West Alameda Avenue, Burbank,
California 91505-4455, United States of America

(72) Inventor
William E Elswick

(74) Agent and/or Address for Service
J A Kemp and Co
14 South Square, Gray's Inn, London, WC1R 5LX,
United Kingdom

(51) INT CL^5
G11B 27/028

(52) UK CL (Edition K)
G5R RB373 RB788 RB81
U1S S2092 S2105 S2116

(56) Documents cited
GB 2224387 A GB 2117168 A GB 2096868 A
GB 2036409 A US 4792794 A
MUSIC MACHINE (The Complete Home Music System) : SPECTRUM USER GUIDE (Amstrad Consumer Electronics PLC and Sinclair Research Ltd.), eg. p3, last para, and page 5, para 2.1(a)

(58) Field of search
UK CL (Edition K) G5R RB81
INT CL^5 G11B 27/028 27/08
On line databases: WPI AND CLAIMS

(54) Digital dialog editor

(57) A system is provided for editing dialog in the digital domain. The system generally consists of three sections, a front end, a plurality of audio processor modules and an input/output system. The front end section interfaces with the user and provides overall system control. In one embodiment, the user interfaces with the system via AT compatible hardware using proprietary software with a mouse (7) and keyboard (8). A graphic representation of the sound is presented on a high resolution graphics monitor (2). An intelligent machine control processor controls the overall operation of the system. Each audio processor module includes a processor for preforming operations on the associated track of data. A shared memory architecture is preferably used, whereby the audio processor modules are linked by a VME bus. In operation, the analog master recordings are converted to digital by the input system. Each track is stored separately for a audio processor module and disk drive. The user may call up and display segments of a graphic representation of sound on the monitor. The sound may be modified by action of the mouse. The system operates to assemble the edited master from the edit decision list.

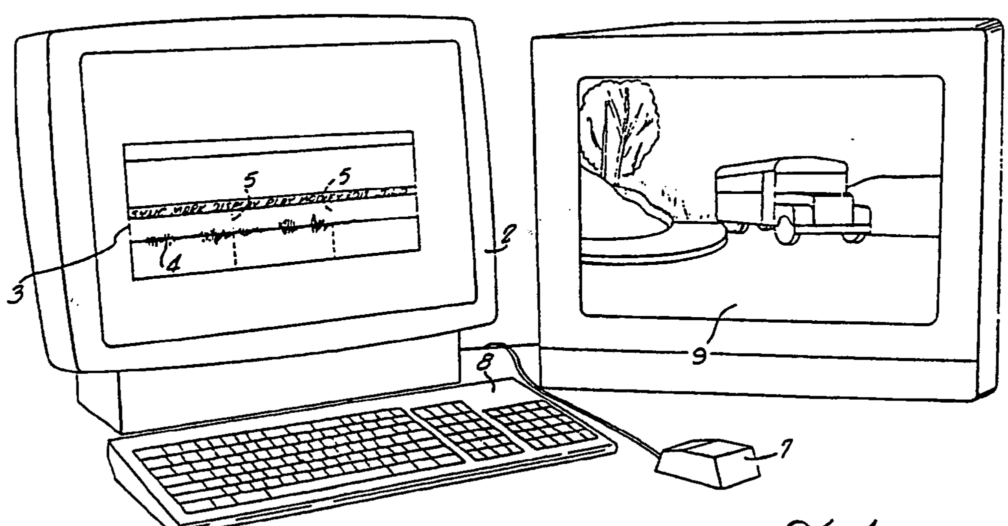


FIG. 1.

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

This print takes account of replacement documents submitted after the date of filing to enable the application to comply with the formal requirements of the Patents Rules 1990.

GB 2 235 815 A

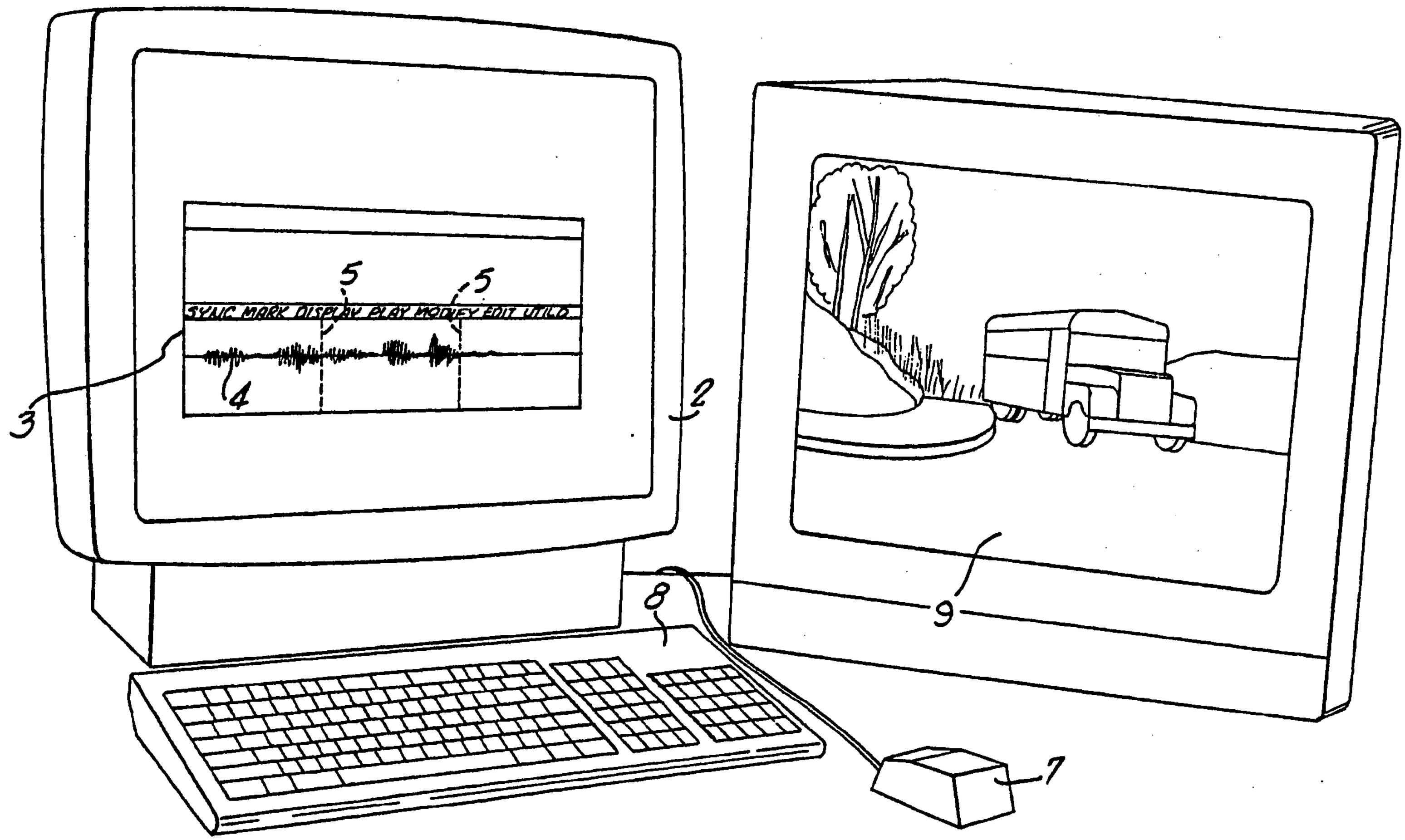


fig.1.

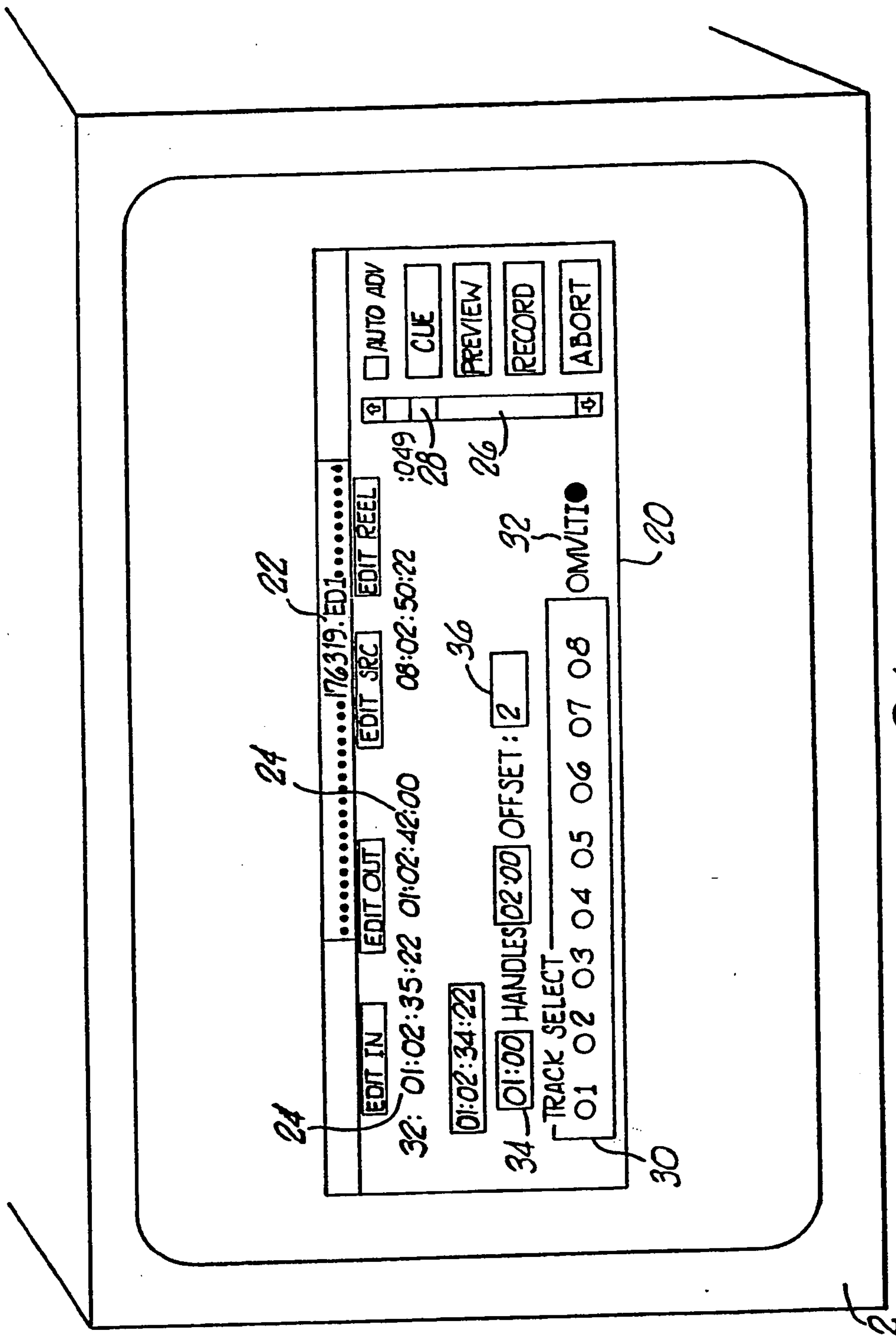

176319.ED1
EDIT IN
32: 01:02:35:22
EDIT OUT
01:02:42:00
EDIT SRC
08:02:50:22
EDIT REEL
:049
AUTO ADV
CUE
PREVIEW
RECORD
ABORT
01:02:34:22
01:00
HANDLES
02:00
OFFSET : 2
TRACK SELECT
01 02 03 04 05 06 07 08
20
22
24
26
28
30
32
34
36
2

FIG.2.

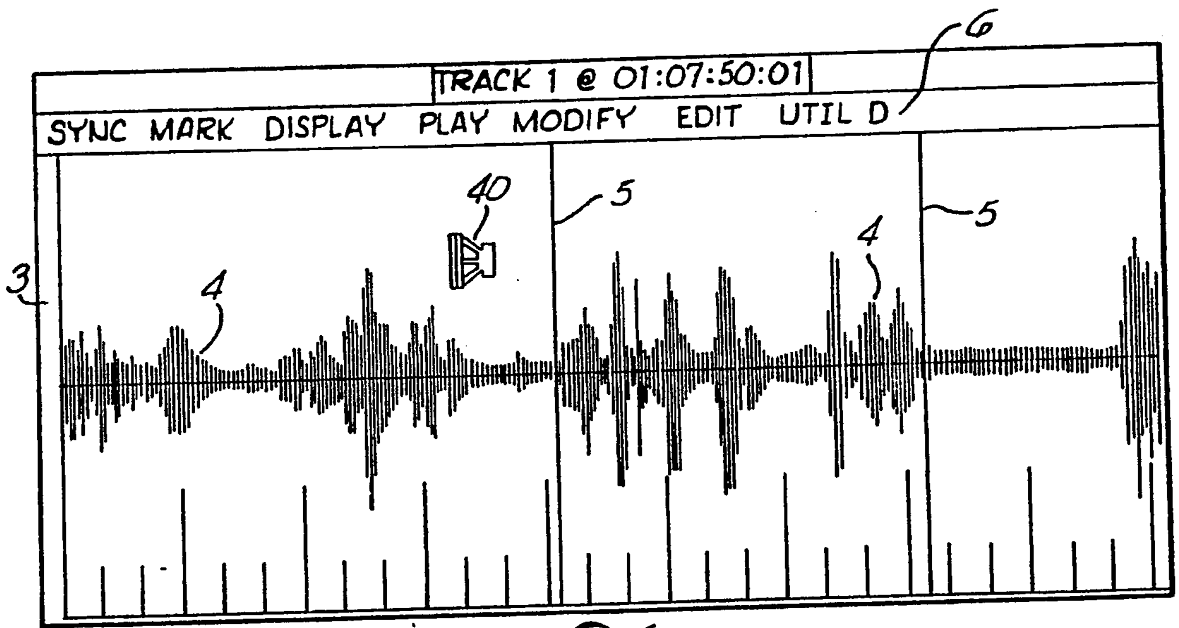

TRACK 1 @ 01:07:50:01
SYNC MARK DISPLAY PLAY MODIFY EDIT UTIL D
6
40
5
5
4
4
3

FIG.3.

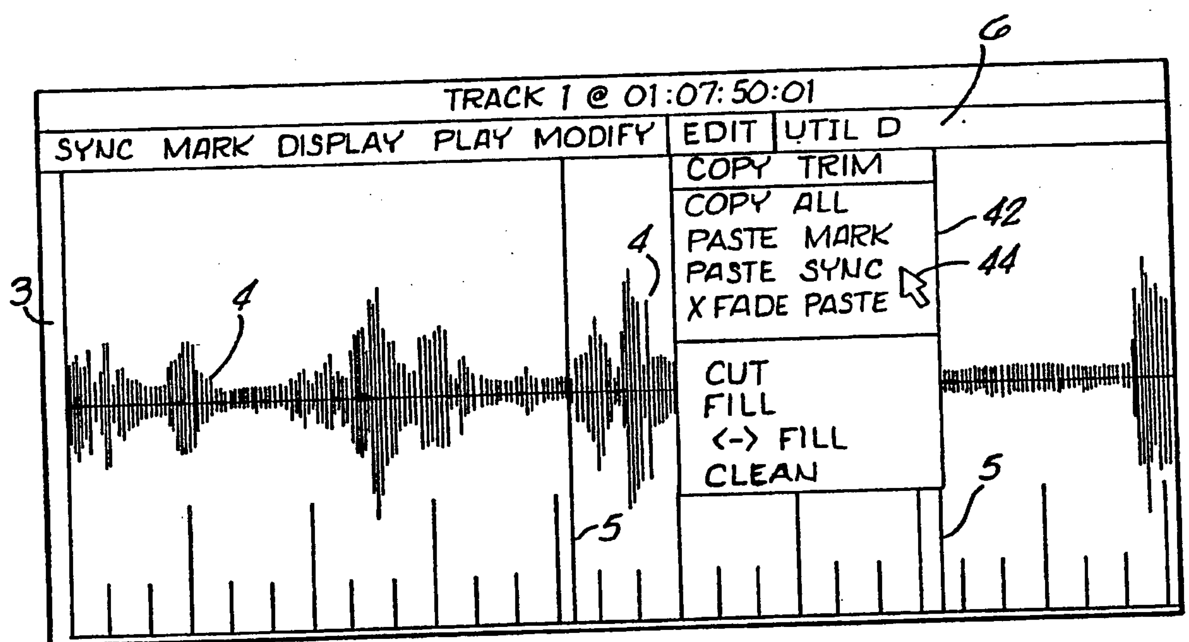
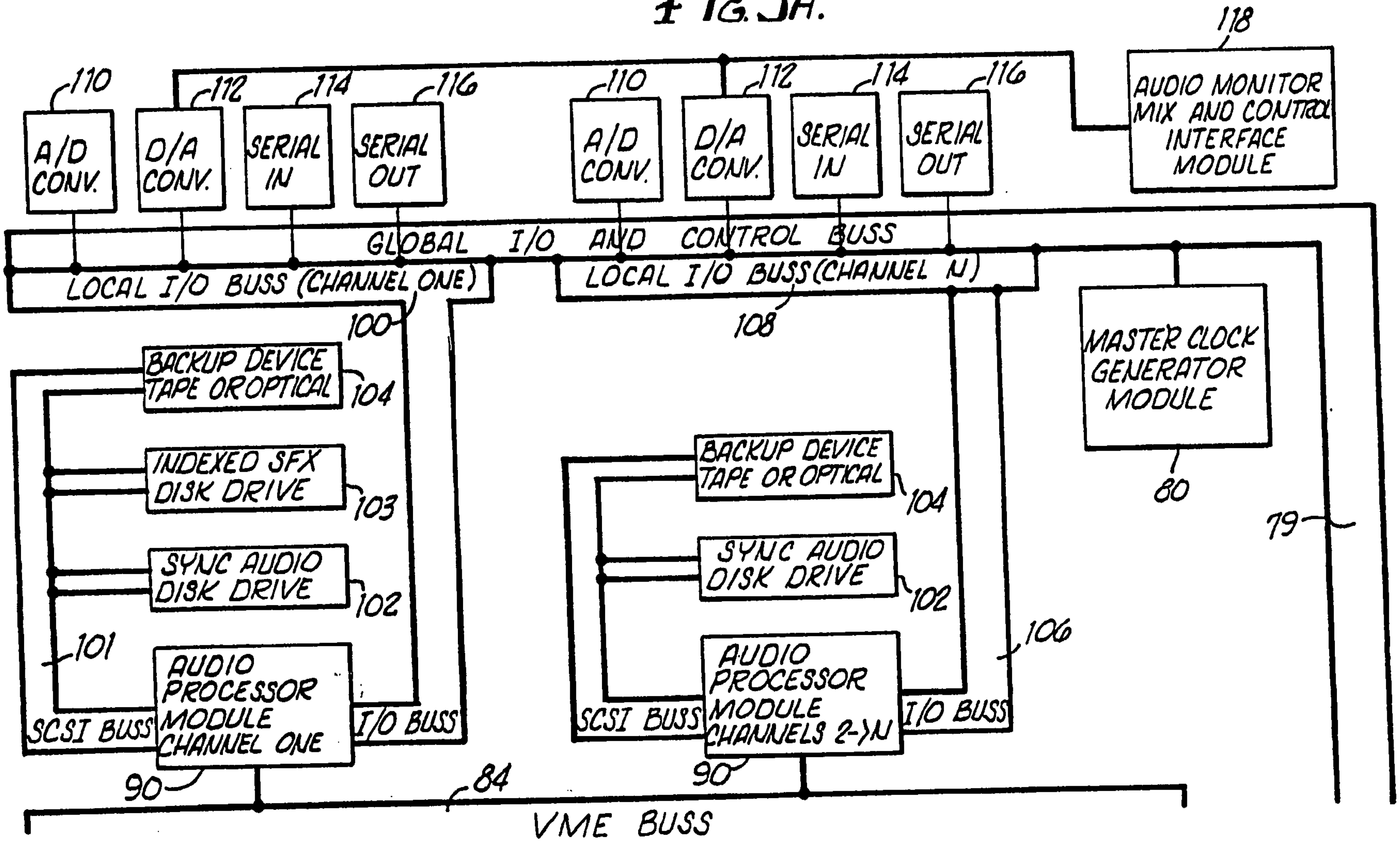

TRACK 1 @ 01:07:50:01
SYNC MARK DISPLAY PLAY MODIFY EDIT UTIL D
COPY TRIM
COPY ALL
PASTE MARK
PASTE SYNC
X FADE PASTE
CUT
FILL
<-> FILL
CLEAN
6
42
44
4
4
3
5
5

FIG.4.

FIG. 5A.


110
A/D CONV.
112
D/A CONV.
114
SERIAL IN
116
SERIAL OUT
110
A/D CONV.
112
D/A CONV.
114
SERIAL IN
116
SERIAL OUT
118
AUDIO MONITOR MIX AND CONTROL INTERFACE MODULE
GLOBAL I/O AND CONTROL BUSS
LOCAL I/O BUSS (CHANNEL ONE)
LOCAL I/O BUSS (CHANNEL N)
100
108
BACKUP DEVICE TAPE OR OPTICAL
104
INDEXED SFX DISK DRIVE
103
SYNC AUDIO DISK DRIVE
102
101
AUDIO PROCESSOR MODULE CHANNEL ONE
SCSI BUSS
I/O BUSS
90
BACKUP DEVICE TAPE OR OPTICAL
104
SYNC AUDIO DISK DRIVE
102
106
AUDIO PROCESSOR MODULE CHANNELS 2->N
SCSI BUSS
I/O BUSS
90
MASTER CLOCK GENERATOR MODULE
80
79
84
VME BUSS

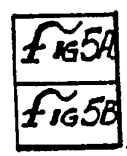


FIG 5A
FIG 5B

FIG. 5.

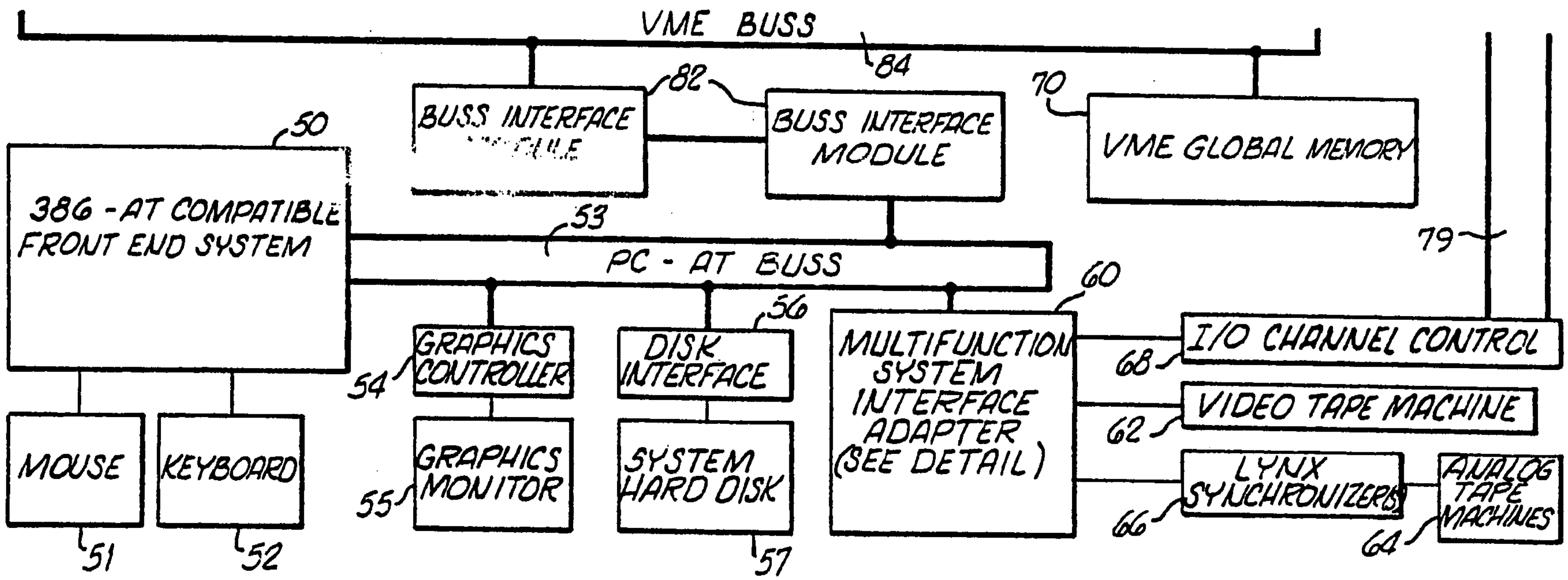

VME BUSS
84
82
BUSS INTERFACE MODULE
BUSS INTERFACE MODULE
70
VME GLOBAL MEMORY
50
386 -AT COMPATIBLE FRONT END SYSTEM
53
PC - AT BUSS
79
60
56
GRAPHICS CONTROLLER
54
DISK INTERFACE
MULTIFUNCTION SYSTEM INTERFACE ADAPTER (SEE DETAIL)
I/O CHANNEL CONTROL
68
VIDEO TAPE MACHINE
62
LYNX SYNCHRONIZERS
ANALOG TAPE MACHINES
66
64
MOUSE
KEYBOARD
GRAPHICS MONITOR
55
SYSTEM HARD DISK
51
52
57

FIG.5B.

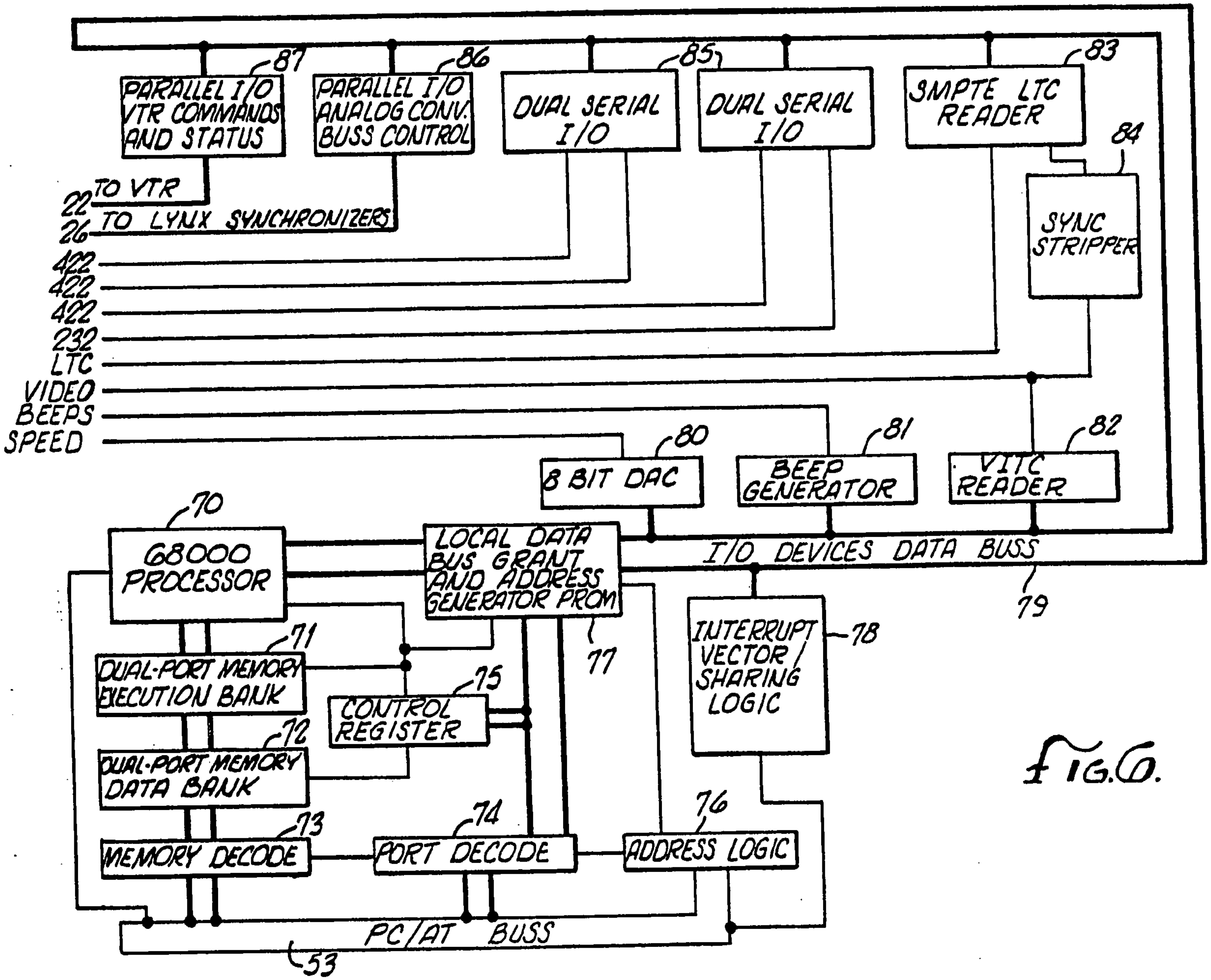

PARALLEL I/O
VTR COMMANDS
AND STATUS
87
PARALLEL I/O
ANALOG CONV.
BUSS CONTROL
86
DUAL SERIAL
I/O
85
DUAL SERIAL
I/O
SMPTE LTC
READER
83
84
SYNC
STRIPPER
22 TO VTR
26 TO LYNX SYNCHRONIZERS
422
422
422
232
LTC
VIDEO
BEEPS
SPEED
8 BIT DAC
80
BEEP
GENERATOR
81
VITC
READER
82
70
68000
PROCESSOR
LOCAL DATA
BUS GRANT
AND ADDRESS
GENERATOR PROM
77
I/O DEVICES DATA BUSS
79
INTERRUPT
VECTOR /
SHARING
LOGIC
78
DUAL-PORT MEMORY
EXECUTION BANK
71
CONTROL
REGISTER
75
DUAL-PORT MEMORY
DATA BANK
72
MEMORY DECODE
73
PORT DECODE
74
ADDRESS LOGIC
76
PC/AT BUSS
53

FIG. 6.

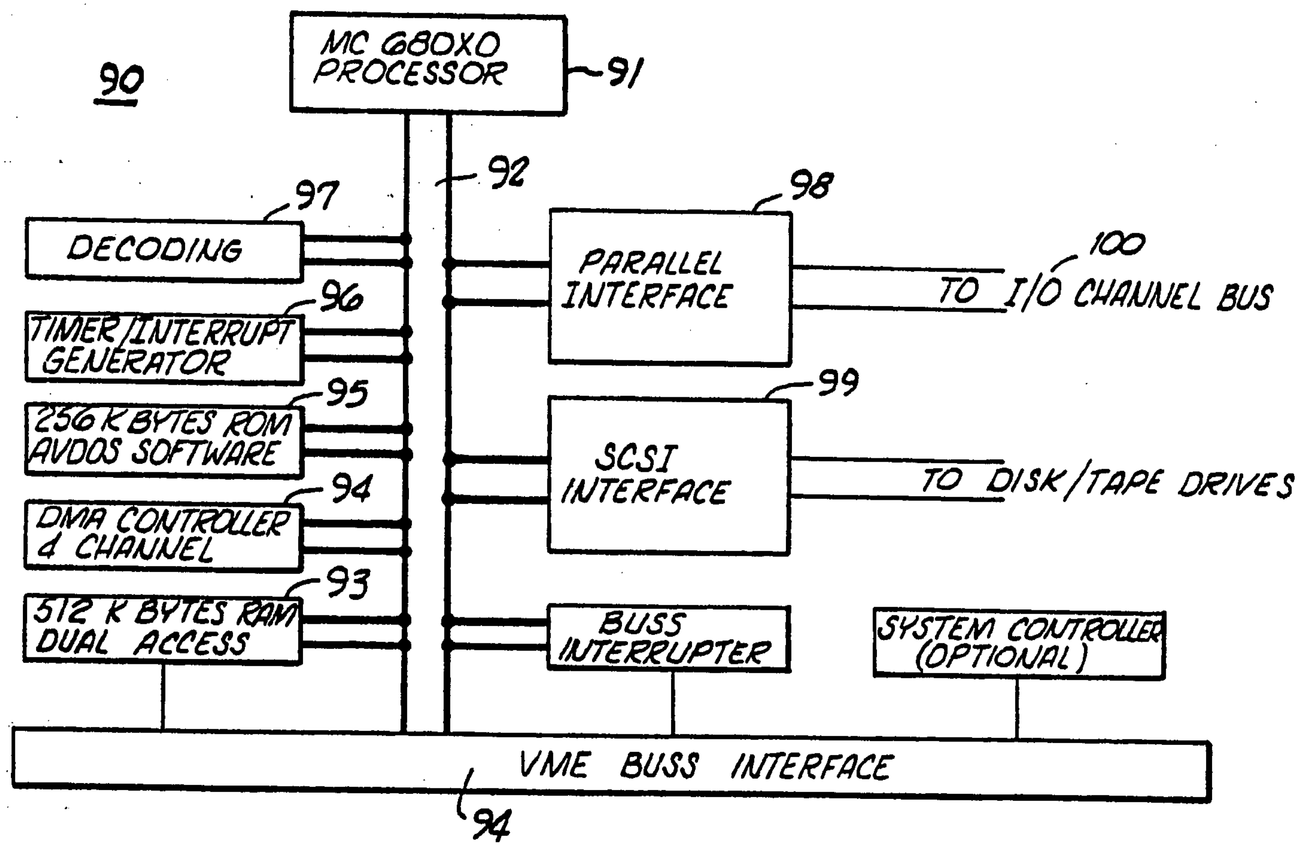

90
MC 680X0 PROCESSOR
91
92
97
DECODING
96
TIMER/INTERRUPT GENERATOR
95
256 K BYTES ROM AVDOS SOFTWARE
94
DMA CONTROLLER 4 CHANNEL
93
512 K BYTES RAM DUAL ACCESS
98
PARALLEL INTERFACE
100
TO I/O CHANNEL BUS
99
SCSI INTERFACE
TO DISK/TAPE DRIVES
BUSS INTERRUPTER
SYSTEM CONTROLLER (OPTIONAL)
VME BUSS INTERFACE
94

FIG. 7.

DIGITAL DIALOG EDITOR

The present invention relates to digital editing for post production film and television production. Specifically, the invention relates to apparatus and methods for digital dialog editing in post production.

The post-production phase of making feature films and episodic television consists of taking the original picture and original sound and converting them into the finished product. The bulk of the work during post-production consists of editing. Editing, in turn, typically consists of arranging the scenes and takes in the final order for the finished product, as it is almost always the case that feature films and television programs are not shot in the same sequence as the finished product. Another major component of editing is modifying the sound, typically consisting of adding sound effects, adjusting

the sound level, mixing the sound with ambiance, adding music and the like. Editing can be extremely time consuming, expensive and difficult. A brief review of the techniques in current use will demonstrate their disadvantages.

During shooting on location or on a stage, the picture and the sound are typically recorded separately. Currently, the picture is most often recorded on film, though recording on video is becoming more common. Sound is ordinarily recorded on analog tape machines, though digital recording is becoming more popular. Thus, the sound is typically recorded on a different media than is the picture, requiring synchronization of the picture with the sound during the post production phase.

For sound recording the most common recording medium is 1/4" audio tape. During production, separate sounds are recorded on separate tracks to achieve the cleanest possible dialog, keeping surrounding noise to a minimum. For example, dialog may be recorded on one track, and ambient production effects recorded on another track.

In order to facilitate resyncronization of the sound with the picture, a reference signal is recorded along with the audio recording during recording. One of the audio tracks records a reference track of pilot tone or timecode, typically SMPTE time code. The SMPTE time code serves to provide an address for particular portions of the audio recording. Machines used during the post-production phase utilize the pilot tone or SMPTE time code to control the speed of the magnetic tape transports

for the audio tape. In this way, a time reference may be maintained throughout the picture making process.

During the shooting of the film or tape, numerous scenes will be shot, often with more than one take per scene. The location sound mixer creates a tape log for each reel recorded. Typically, the tape log includes information such as scene number, take number, the SMPTE time code, as well as additional comments. Commonly, the takes to be used in the picture cut, known as master takes, are decided on location. These are noted on the tape log by circling the take number.

In a process known as off-line editing, a rough cut version of the final film is generated. The selected master takes are transferred from the sound reels in synchronization with the picture, to either a video medium for tape, or a film print called a work print for film. The SMPTE reference code for each cut is logged to create a list of scenes, called an edit decision list. The edit decision list may be compiled manually or by computer.

In the rough cut version, both the sound and picture are first generation copies of the original picture and audio recording. With analog sound recording techniques, sound suffers a loss of frequency response, the relative loss of high and low frequency sounds, as well as an increase in the tape noise, high frequency hiss, with every generational copy.

Next, in a process known as on-line editing, the final version of the film or tape must be constructed from the original film or tape, and ideally, using the sound from the

reels recorded on location. Here, two main actions take place. First, the original picture and audio must be rearranged from the order in which the scenes and takes were originally shot to that specified by the edit decision list.

Second, the audio track must be evaluated for modification or improvement. Ordinarily, in a process known as spotting, the rough cut audio track is evaluated for editing. Dialog may be spotted for automated dialog replacement (ADR), a process in which dialog is re-recorded in synchronism with the picture to replace the original production dialog. In this way, excessive background noise, such as planes or automobiles, found in the original recording, may be eliminated. Further, it may be desireable to have the line delivered differently, or indeed, to change the line. Another major use for spotting is in a process called foley (the replacement of production sounds created by humans or props, such as footsteps, body movement or paper rustling or paint brush movement) which are commonly low in level and require re-recording or replacement with stock sounds from a library.

In a process known as prelay, multiple tracks of sound are generated. Commonly, the audio is put onto different tracks, for example, dialog being on several tracks, sound effects on other tracks, foley on yet other tracks and music on still other tracks. It is not uncommon to have upwards of twenty tracks by the time of prelay. The tracks are then mixed in the desired proportion in the final mix. Of all of the components in the prelay, the dialog is the sound element which is the most

predominant in the final mix. Accordingly, it is typically the least forgiving of flaws.

Overview of the Sound Editing Techniques of the Prior Art

There are three potential recording mediums on which to build soundtracks. The first, and oldest, recording medium is film coated with a magnetic medium. The traditional editing method with magnetic film is to physically cut and splice magnetic film. The second medium uses analog tape. This method originated from the sound recording industry. The analog tape method involves recording on multi-track tape. The third medium uses electronic or optical storage, such as recording digital audio to a hard disk drive or other random-access digital storage media.

1. MAGNETIC FILM -- Sound editing on magnetically coated film is the traditional method used in the feature film industry. First, the selected scenes and takes are transferred from the original recording on 1/4" magnetic audio tape to magnetically coated film stock. The coated film stock recording is one generation down from the original master. The coated film stock is then physically cut, the sections being spliced and pasted together manually, much as a linear collage. The audio and film are literally physically synchronized.

An editor generates physically separate reels of film for the various sound tracks. Thus it is not uncommon to have anywhere from three to twenty reels of audio film at a single time. Dialog may be assembled on one reel, sound effects on

another and music on yet another. If one track has a period in which no sound is present, blank magnetic film stock is inserted for that duration.

To achieve sound level changes the editor physically scraps the tape with a razor blade or uses a solvent to remove some of the oxide coating. Substantial skill and experience is required to achieve the right sound level without damaging the overall sound quality.

Typically, editors build their tracks of audio on film work benches. A work bench consists of a sprocketed box with reel holders on both sides. The boxes have positions for one picture reel and three audio reels. The workbench has various bins in which the editor may temporarily store sections of recorded magnetic film, to avoid misplacing or confusion of film trims. The editor then removes the film trims from the bins and assembles them in the order set by edit decision list by splicing them at the appropriate position in the given reel.

Generally, the advantages of the film cutting technique are that minimal equipment is necessary, and that last minute changes are relatively easily accomplished by insertion of additional film stock. Further, by the arrangement of the editors workbench, the trims to be assembled may be physically arranged, thereby making for a conceptually straightforward task in assembling the trims. Generally the disadvantages of the film cutting technique is the lack of any reasonable way to process sound. Scraping the oxide and use of solvents requires the use of extremely skilled labor and is time consuming.

Further since the magnetically coated film is one generation removed from the original master, the sound quality is degraded relative to that of the master. Finally, the resolution in synchronizing the magnetic film to the picture film is limited to a resolution of one sprocket hole.

2. MAGNETIC TAPE TECHNOLOGY -- Editing on tape utilizes equipment originally developed for the music industry. Specifically, multiple track recorders can record up to 24 separate channels on a single 2" magnetic audio tape. While the recording format on the tape may be either analog or digital, digital recording has not been used often for editing on tape typically because of the higher cost for transports and format incompatibility.

Typically, the audio recording from the location or stage comes with the SMPTE time code reference recorded. A channel on a multiple track tape is provided with the SMPTE time code reference. Since the SMPTE time code reference is used to index the audio to the film, the audio tapes may not be cut and spliced as was done in the case of the magnetic tape technique. With magnetic tape, the take to be transferred from the original audio tape is set to a few seconds of tape before the take to be transferred. Similarly, the multi-track audio tape is positioned slightly ahead of the position where the transfer will occur. Then, both the daily and the multi-track tape machines are brought up to speed and synchronized using the SMPTE timecode. At the designated SMPTE time code the audio

information from the daily is transferred to the selected track of the multi-track tape.

The advantages of the magnetic tape approach over film is that the sound quality reproduction of the multi track is generally superior to that of magnetic film. Magnetic tape has a wider frequency response and can handle louder sound levels without distortion. This in turn decreases the amount of audible tape hiss. Further, the edit can be assembled from an edit decision list by computer control.

The disadvantage of magnetic tape relative to film is that changes after the prelay, changes known as conforming, are extremely difficult to perform. With audio tape it is not possible to physically cut the tape containing the time code and insert blank tape for delay. Conforming with audio tape requires a generation loss by re-recording, unless a digital transport is used. Further, editing dialog on audio tape often requires multiple attempts to achieve a workable edit, since the editor must perform some operations such as fades in a live performance during the edit. Each attempt requires rolling the tape machines back to a point before the edit and resynchronizing the machines before the edit can be attempted. Finally, since the tape machines must be rolling and in synchronism with each other to perform an edit, the necessary rolling and previewing is time consuming and contributes to faster deterioration of audio quality due to tape wear. Finally, when using an automated assembly from the edit decision

list, the resolution accuracy is limited to the time difference between time codes, which is typically one frame.

3. OTHER DIGITAL AND HARD DISK SYSTEMS -- Recently, it has become possible to record sound in a digital format. The benefits of digital recording are generally that the sound reproduction quality is better than magnetic film or analog tape. Generally, unlike film or tape, a digital disk can have several backup safety copies with no generation loss. Further, deterioration in long term storage does not occur. Finally, the space required for storage is considerably less than for film or for tape.

Attempts have been made to use digital systems designed for musical applications in dialog editing. These applications range from the use of sampling keyboard synthesizers in conjunction with tape systems to various hard disk recording systems developed for the music industry. These attempted solutions have not been successful when applied to dialog editing, generally because the machines are designed to solve a different set of problems than those encountered with dialog editing.

Attempts have been made to utilize digital techniques for sound effects in dialog editing. Sound effects are typically built from short segments of sound. Digital samplers designed to sample single notes from acoustic or electric instruments have been used successfully with sound effects. For dialog, however, these sampling systems have serious shortcomings. A typical sequence of dialog will utilize far more memory than the

maximum capacity of even the largest available samplers. To use a sampler, the dialog would have to be repeatedly transferred back and forth from a separate multi-track recording device in order to form an edit. On anything less than a giant sampler, this approach is impractical.

SUMMARY OF THE INVENTION

Post-production of feature films and episodic television is performed with digital audio. The system utilizes the best aspects of old film based and magnetic tape editing, but successfully utilizes the advantages of digital audio. The highest quality digital audio edit is achieved at reduced cost in minimal time.

The editing paradigms are abstracted from methods used by working editors. Short segments of sound, analogous to short segments of film or 'trims' in manual editing, are converted to visual displays on a screen. The editor modifies the on-screen trim by operation of a mouse and keyboard. Typical edits might include selectively editing out a word or sound, performing fades, modifying the volume or copying sounds. The edited trim may be stored in memory while still being displayed, analogous to the use of 'bins' in manual editing. Since the sound to be edited is stored in memory, full digital quality is maintained. After editing, the trim may be restored to disk.

After the various trims have been edited, the complete sound track is assembled. The system uses the edit decision

list to automatically generate the digital edited master. This permits rapid assembly of the master.

The hardware of the system generally consists of three sections -- a front end, a plurality of audio processor modules and an input/output system. First, the front end section interfaces with the user and provides overall system control. In one embodiment, the user interfaces with the system via AT compatible hardware using MicroSoft Windows with a mouse and keyboard. A graphic representation of the sound is presented on a high resolution graphics monitor. An intelligent machine control processor controls the overall operation of the system.

Second, each audio processor module includes a processor for preforming operations on the associated track of data. Each audio processor module has a disk drive associated with it for mass data storage. A shared memory architecture is preferably used, whereby the audio processor modules are linked by a VME bus. Third, the input/output system may include analog to digital and digital to analog converters, providing an interface between a typically analog world and a digital editing system.

In operation, the analog master recordings are converted to digital by the input/output system. The assembly of the various takes may be done under computer control based on the edit decision list. Each track is stored separately on a disk drive associated with an audio processor module. If editing of individual 'trims' is desired, the editor may call up and

display segments of a graphic representation of sound on the monitor. The sound may be modified by action of the mouse. After the individual edits are preformed on the 'trims', the system operates to assemble the edited master.

Accordingly, it is an object of this invention to provide a system capable of performing editing of digital audio, particularly editing dialog as digital audio.

It is a further object of this invention to provide an editing system which uses editing paradigms abstracted from those used by working editors, thereby creating a virtual 'editor's work bench' from digital electronics.

It is another object of this invention to provide a graphic representation of sound to permit editing of that sound.

It is yet a further object of this invention to provide auto-assembly of the edited digital material from an edit decision list.

It is a further object of this invention to provide a hardware architecture which promotes rapid digital editing.

It is a further objective of this invention to provide a system clock which is synchronized to a video signal.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a perspective view of the user interface, including a screen display with a typical edit window.

Fig. 2 is a representation of the screen display for the assembling process.

Fig. 3 is a representation of the screen display and of an edit window.

Fig. 4 is a representation of the screen display and of an edit window with a pull-down display.

Fig. 5 is a block diagram of the digital dialog editor system.

Fig. 6 is a block diagram of the intelligent machine control interface.

Fig. 7 is a block diagram of the audio processor module.

DETAILED DESCRIPTION

Overview Of The User Interface

Referring to Fig. 1, a typical user interface display is shown. Specifically, on a monitor 2 there will be a window 3 in which a graphic representation of sound 4 is displayed. Typically the representation of sound 4 is depicted as audio oscillation as a function of time. Within the window 3, the three fairly localized sections might be, by way of example, three separate words of dialog or three occurrences of sound effects. The 'trim', or segment of sound on which the editing operation will be performed, may be the whole window 3 or some smaller portion, as defined for example by vertical marks 5. Because of the presentation environment chosen for the preferred embodiment, the menu selections 6 are displayed continuously on the screen. User input and selection is made by input devices, by way of example, a mouse 7 or keyboard 8.

Typically, a video monitor 9 is used to display the picture which accompanies the sound. Ultimately, exact synchronization of the picture and sound is done by the audio and visual comparison by the editor.

The Assembly and Editing Process

1. The Assembly Process -- The first step in the dialog editing process is the selection of takes to be assembled. Typically, an edit decision list is generated during the off-line editing process. This results in a listing of the takes to be assembled, often specified by reel number and SMPTE timecode numbers of the starting and stopping locations. In the absence of an edit decision list, the user may choose to create one during spotting, or enter the data through the window on a take by take basis.

The second step in the assembly process is to transfer the sound from the original recording medium to the digital dialog editing system. This process is controlled by the user via the assembly window. Fig. 2 shows a typical representation of the assembly window. Broadly speaking, the assembly process consists of the steps of: first, determining what take is to be transferred from the original recording medium to the digital dialog editing system, second, causing that transfer of the sound, and third, controlling the receipt of the sound and storing it on the electronic media in the digital dialog editing system.

Referring to Fig. 2, the assembly window 20 controls the capture of sound takes from tape to disk. In determining what take is to be transferred from the original recording medium to the digital dialog editing system, the edit decision list may be used if one is available. The user may load the edit decision list into the window, and an identification of the list 22 may optionally appear on the screen 2. The starting and stopping times are displayed 24. In the preferred embodiment, a time bar 26 is used to increase or decrease the time. The cursor 28 is positioned by movement of the mouse 7 (Fig. 1) and a clicker selects the action, for example, time increase or time decrease.

Whenever possible, an editor will use the take that corresponds to the shot chosen by the picture editor. If the picture edit decisions have been made on an electronic system, the edit decision list can be used directly to locate sound takes and picture edits. The assembly window prompts the user to load the necessary sound reel for the specified take and locates the take using timecode. To load a take, the user selects a track and the audio is recorded digitally. Any or all of the tracks may be selected by activating the track select locations 30. Recording may be done to all tracks by selection of the multi 32 option. At this point, the video, disks, and source machine may be synchronized.

During the assembly process, it is often desirable to add handles, that is, transferring a few seconds of sound before and after the actual sound for the edit. The assembly window may be used to automatically add handles, with the desired handle time

displayed 34. Additionally, it is often necessary to offset the recording as the sound is transferred. The desired offset may be set and displayed 36.

The edit may be simultaneously displayed in the corresponding track window in case any adjustment is necessary. The editor may choose to make corrections at this point or continue directly to assemble the next take.

2. The Editing Process -- Sound takes are not always usable in their original form. The recording may contain dolly squeaks, generator noise, alien footsteps, or production sound effects that should be kept, but not on the dialog tracks. In addition, lines that will be replaced in ADR are usually kept for comparison. Typically, the editor will keep anything that might be usable in the mix, but split sounds to separate tracks in such a way as to maximize the choices of the mixer. Offending sounds that must be kept can be moved to extra tracks set aside for the purpose. Typically an editor will dedicate tracks for production effects, one or more tracks of alternate dialog, and dialog that will be reprocessed.

Once the sounds are split to separate tracks, the main dialog tracks must be filled in such a way that the 'air', or background ambiance, will play smoothly behind the other tracks. Since lines replaced by ADR will be recorded in a soundproof booth, the dialog tracks must supply the ambience to make it seem that the actor delivered the lines on the set. This ambience may consist of body movement, air movement, or any sounds native to the environment in which the scene is set.

To clean up a take, the editor usually finds a piece of fill, that is, clean background noise, with which to replace unwanted noises on a dialog track. On a system as quiet as the digital dialog editor, this may even be tape hiss if the scene was recorded in a very quiet environment. The fill can be repeated or 'looped' to form a longer section than whatever is available in the original recording, and saved in a window or the bins for later use.

The digital dialog editor may provide editing operations to support these operations through the concepts of trims and timecode editing. A trim is any section of sound the user can see in a window. A trim is defined by marks 5 in the window or by the window 3 itself if no marks are set. Editing can also be performed by specifying timecode locations. Typical operations would include the functions such as copy, erase, and move. Since sound really does not exist in frames the way picture does, only relatively gross edits can be performed with the desired degree of precision using timecode numbers, which only specify frames. Trims, however, may be specified down to a particular digital sample, typically on the order of one 40,000th of a second, if necessary.

Fig. 3 shows a display of a typical edit window. A visual display of the sound 4 is provided. By moving an icon 40, represented as a speaker in the preferred embodiment, the sound may be played via a speaker (not shown) to the user. By moving the icon 40 in a horizontal direction relative to the displayed sound 4, the sound may be played. The sound is played in the

speed and direction the icon 40 is moved. This function is analogous to the so-called scrubbing function of the prior art, where an audio tape reel is rocked back and forth over the pick-up heads to hear the sounds. In this way, the user determines the exact location of the sound to be edited. Through visual inspection of sound in windows and the use of a scrubbing function, edit points can be specified without typing numbers to the necessary resolution. By varying the display resolution, the user can zoom in on areas of interest to enhance accuracy, or zoom out to provide context.

Once a trim is selected in a window 3, the user may select from one of several menus 6 the operation to be performed. Operations are grouped in menus 6 by the class of operation involved. If an operation involves more than one window, the operation is selected in the destination window. Fig. 4 shows a typical display for which the EDIT operations 42 have been pulled down. The cursor 44 is moved by action of the mouse for selection of the desired function.

Any of the various edit functions which are performed in editing magnetic film or magnetic tape may be implemented here. Generally, any sound level adjustment may be made by multiplying the stored sound amplitude by an adjustment factor. For example, the sound in an area of a window 3 marked by mark 5 may be multiplied by a reduction or increasing factor. This is called a digital 'shave', since it is similar to shaving of oxide from magnetic film. A digital 'scrape' may be performed, where the amplitude drops fairly rapidly, and then recovers to

the original level fairly slowly, or conversely, drops fairly slowly and rises fairly rapidly. A fade out maybe done by marking the head and tail mark 5 and selecting the fade out option. The fade out may be done linearly, logarithemically or parabolically with respect to time, or in any desired manner and speed. The blend function adds two tracks together, and permits amplitude modification of the source tracks. The 'sync slide' function moves a marked section of sound and provides a fill sound in the marked section. The 'sync paste' function copies from one track to another. The 'cross fade/paste' function provides variable cross-fading from a source, which is added to a destination track.

Overall Architecture, Hardware and Software.

For most applications, implementation of the digital dialog editor system requires considerable processing power in the hardware. Ideally, each part of the system must perform tasks quickly and efficiently, and in close cooperation with other parts of the system. Sub-systems may be optimized for their respective tasks and linked via a shared memory architecture. In the preferred embodiment, each sub-system contains private memory as well to minimize contention for resources and allow efficient multi-tasking across multiple processors.

Conceptually, the digital dialog editor can be broken down into four major sections. These are the 'front end' or user interface, the realtime control section, the audio operating system, and the audio processing system.

1. The User Interface -- The user interface portion of the system may be understood with reference to Fig. 5. The user interfaces with the system through processor or computer 50. In the preferred embodiment, the choice of Microsoft Windows as the presentation environment dictated the choice of Intel family microprocessors for the user interface. An Intel 80386 microprocessor based system was chosen for its superior speed and memory management capabilities in a multi-tasking environment. It will be appreciated by those skilled in the art that any computer or processing system with the appropriate functionality, such as a minicomputer or Apple Computer, may be employed in place of the 80386-AT compatible system employed in the preferred embodiment. The user may interface with the computer 50 by the mouse 51 and keyboard 52. A graphics monitor 55 is used to display the graphic representation of the sound. In the preferred embodiment, the display is generated using an enhanced VGA controller 54 which provides 16 colors and a resolution of 1024 by 768 pixels. Memory, preferably in the form of a hard disk 57, are provided for the computer 50. Access to the hard disk 57 and graphics monitor 55 are made from the computer 50 over the PC-AT bus 53 via the disk interface 56 and the graphics controller 54, respectively.

The front end includes that which is visible to the user of the system. The display interface code is provided by the Microsoft Windows operating environment. Since this environment provides preemptive multitasking support, several processes can appear to be running simultaneously. In addition to visualizing

the process for the user, the front end code checks commands for errors, provides prompting for complex operations, and monitors background processes the user may choose to invoke.

2. The Realtime Control Section -- The intelligent machine control block 60 interfaces with the computer 50 over the PC-AT bus 53. The intelligent machine control/interface processor 60 controls a wide assortment of external machinery, such as a video tape machine 62, one or more analog tape machines 64 through Lynx synchronizers 66 and input/output channel control circuitry 68. All of the real time processes required for external machine control may be handled by the machine control circuit 60, allowing the other sub-systems to operate external devices with almost no overhead.

The realtime control code supervises and coordinates the operations of the various parts of the system. It handles location and synchronization of tape machines and the audio processing system. In the preferred embodiment this code is really split across two processors: one part runs at the MS-DOS system level on the front end machine, the other part runs on the intelligent machine control card. The code on the machine control card contains another multitasker so that multiple machines can be controlled simultaneously.

Fig. 6 provides detail as to the structure of the intelligent machine/interface processor 60. In the preferred embodiment, a Motorola 68000 series microprocessors 70 is employed. The microprocessor 70 is connected to the bus 53 via

bank 72 and memory decode circuit 73. When the machine control 60 is to provide selection of the input/output channel, information transferred over the PC/AT bus 53 is taken in through the port decode 70 and/or utilized by the local data bus grant and address generator PROM 77. Control register 75 may receive the output of the port decode circuitry 74 and provide information to the microprocessor 70, dual port memory execution bank 71 and dual port memory data bank 72.

In the preferred embodiment, the machine controller 60 includes a SMPTE time code reader 83, and four serial interfaces 85, which may be of the RS-232 or RS-422 interface. Additionally, two parallel machine interfaces are provided, one for VTR commands and status 87 and one for parallel input/output analog converter bus control 86 for the analog tape machines 64. Each interface may be programmed to support a particular interface protocol independently of the others. This permits the supervising software to control any type of machine without concern for the details of a particular machines characteristics. In this way, even machines with rudimentary capabilities may be controlled in a sophisticated manner.

Returning to Fig. 5, the input/output device data bus 79 is shown interfacing with the master clock generator clock module 80. The master clock generator clock module 80 serves to lock the system audio sampling rate to a video signal. In this way, an audio sampling frequency which is a known fraction of a well established video frequency is generated. The master clock frequency for color video in the NTSC system is 14.31818 MHz.

The master clock frequency is generated by frequency multiplying by a factor of 4 the color burst signal (3.579545 MHz). The color burst signal is extracted from a common composite reference known as color black. The audio sampling rate is 44055.94 KHz, which is the master clock frequency divided by 325. The master clock generator module 60 supplies the master sample clock to each channel of the system. It provides a means to lock to an external video reference source or external sample clock for precise synchronization.

3. The Audio Operating and Processing Systems -- The system contains a plurality of audio processor modules 90. In Fig. 5, the left most audio processor module 90 is labeled channel one, and to the right is a single representation intended to be repeated for the various channels.

In the preferred embodiment, each channel processor runs it's own copy of the audio operating system. The audio operating system provides interface services for commands and devices to the audio processing software. Details like buffer management are accounted for in this section. It also supplies display lists and status data to the front end.

Referring to Fig. 7, the details of an audio processor module 90 are shown. In the preferred embodiment, a Motorola 68010 (or 68020) microprocessor 91 controls other special purpose devices for input/output and data manipulation. A bus 92 connects microprocessor 91 to the VME bus 94, as well to a dual access RAM which spans the VME bus 94 and audio processor bus 92, a four channel DMA controller 94, memory 95 for storing

the program for the microprocessor 91, a timer/interrupt 96 and a decoder 97. The microprocessor bus 92 provides output to the local input/output bus channel 100 via the parallel interface 98. The audio processor module 90 is connected to a variety of storage devices, including the synchronized audio disk drive 102, backup device 104, and an indexed sound effect disk drive 103 via a small computer system interface (SCSI) bus 101.

In the preferred embodiment, each channel also runs a copy of the audio processing software. This consists of the fundamental editing commands such as blend, fade, scrape, copy, etc. Although this section may contain only very simple commands, many complex editing commands may be located here for the sake of speed.

Since the dialog editor can edit sound in memory as well as on disk, VME global memory 70 is provided which may be accessed over the VME bus 84. In the preferred embodiment, 8-12 megabytes of RAM are shared by the audio processor modules 90 for use as an edit buffer. The memory 70 is allocated to a particular processor on request.

In the preferred embodiment, each channel has a dedicated disk 102 for mass data storage. In a typical configuration, all but the last channel uses a Winchester hard disk drive. Use of the SCSI interface allows easy configuration of systems with various sizes and types of storage devices. Hard disk drives of up to 760 megabyte capacity are commercially available.

In the preferred embodiment, at least one channel of each system is provided with a disk drive using magneto optical

technology. The media for these drives is removable and reusable. While these drives are not as fast as the fastest Winchester drives, they are fast enough for real time digital audio applications. Each optical disk cartridge provides 600 megabytes of data storage, 300 megabytes per side. When not in use as a channel disk, the magneto optical drive serves as a backup device for the other channels.

A method termed virtual disk sectoring is used in which the rudimentary operations of the disk operating system present the structure of the disk drive as some number of logical blocks or sectors of arbitrary size. At the time that the system loads, a boot block is read from the first physical sector of the disk drive 103 or 104 which contains, among other things, the desired virtual sector size. All subsequent disk operations are specified in logical (virtual) sectors based on this virtual sector size. The virtual size is adjusted so that there is a one-to-one correspondence between the time duration of one virtual sector's worth of digitized audio and the time duration of rate of the audio and the frame rate of the picture. A typical example would be Video at a frame rate of 29.97 frames/sec and audio sampled at 44055.9 samples/sec. In this case, the size of a virtual sector would be 1470 samples (16 bit words) or 2940 bytes.

This relationship can be expressed generally as:

$$S(f) = S(s) * (1 / F(s))$$

Where S(d) = samples/frame;

S(s) = samples/sec;

F(s) = Frames/second

and for given frame rate, a sample rate is chosen such that S(f) is an integer. This method gives a direct correspondence between frames of picture and virtual frames of audio.

In a further effort to reduce access time to the mass storage devices 102, the synchronized digital audio is stored sequentially on the disk drive 102. That is, as the edited sound is assembled, it is stored on the disk in sequential, rather than random, fashion. In this way, the fully edited audio may be read from the disk drive 102 in real time, without large gaps of time which would be incurred in accessing the disk drive 102 randomly.

4. The Input/Output System.

Because most audio for feature films and episodic television is recorded in analog, analog to digital converters must be used to generate the digital audio. However, should the original recording be done digitally, no further conversion is necessary. Referring to Fig. 5, separate analog to digital converters 110 are preferably provided for each channel. Similarly, separate digital to analog converters 112 are provided for output from each channel. The converters 110 and 112 each connect in parallel to the local input/output busses 100 and 108. Optionally, serial input devices 114 and serial

output devices 116 may be provided. The output of the digital to analog converters 112 is provided to an audio monitor and mix and control module 118, of any of the types known to those skilled in the art.

Though the invention has been described with respect to specific preferred embodiments thereof, many variations and modifications will become apparent to those skilled in the art. It is therefore the intention that the appended claims be interpreted as broadly as possible in view of the prior art to include all such variations and modifications.

What is claimed is:

1. An editing system for editing multiple channels of sound, comprising:

an input including an analog to digital converter for each channel,

an audio processor module for each channel,

a disk drive for each channel,

a bus system interconnecting the audio processor modules for each channel,

a user interface,

a machine control system, and

an output.

2. The editing system of claim 1 wherein the user interface provides a graphical display of the sound during editing.

3. The editing system of claim 2 wherein the user interface provides a graphical display of an icon.

4. The editing system of claim 3 wherein the icon is a graphical depiction of a speaker.

5. The editing system of claim 2 wherein the user interface includes mark lines for identifying the sound to be edited.

6. The editing system of claim 1 wherein the user interface is a general purpose computer.

7. The editing system of claim 1 wherein the user interface further includes a mouse.

8. A system for editing sound comprising:

an input system for receiving the sound to be edited,

an audio processing module,

a user interface including a graphic display of the sound, and

an output system for outputting edited sound.

9. A method for editing sound in a memory based editing system comprising the steps of:

storing in memory the sound to be edited,

displaying a graphic representation of the sound to be edited,

indicating edits to be performed on the graphic representation of the sound,

editing the sound in accordance with the edit indicated on the graphic representation, and

storing the edited sound in memory.

10. The method for editing sound of claim 9 wherein the edit consists of a digital scrape.

11. The method for editing sound of claim 9 wherein the edit consists of a digital shave.

12. The method for editing sound of claim 9 wherein the edit consists of a fade out.

13. The method for editing sound of claim 9 wherein the edit consists of a local area mixing.

14. The method for editing sound of claim 9 wherein the edit consists of a spot level adjustment.

15. The method for editing sound of claim 9 wherein the edit consists of a sync paste.

16. The method for editing sound of claim 9 wherein the edit consists of a crossfade paste.

17. The method for editing sound of claim 9 wherein the edit consists of a sync slide.

18. The method for editing sound of claim 9 wherein the step of storing the sound to be edited to memory includes the steps of controlling the machines from which the sound is transferred, and controlling the storage of the sound in the memory.

19. The method for editing sound of claim 9 wherein the step of storing the sound to be edited to memory is determined based upon an edit decision list.

20. The method for editing sound of claim 9 wherein the step of performing edits on the graphic representation of the sound includes the steps of marking the area in which the edit is to be performed and selecting the editing function to be performed on the marked area.

21. The method for editing sound of claim 9 wherein the step of performing edits on the graphic representation of the sound includes the step of causing an icon to move relative to the graphic representation to perform a scrubbing function.

22. A method for generating a digital audio sampling rate for use with a system having a video frequency comprising the steps of:

dividing the video frequency by a set amount, and

setting the divided frequency as the digital audio sampling rate.

23. The method for generating a digital audio sampling rate of claim 22 wherein the video frequency is the master clock frequency for color video in the NTSC system.

24. The method for generating a digital audio sampling rate of claim 22 where the set amount by which the video frequency is divided is 325.

25. A method for editing sound comprising the steps of:

displaying in an edit window a graphic representation of the sound to be edited,

indicating which section of the graphic representation in the edit window is to be edited and

selecting the edit function to be performed from a menu.

26. The method for editing sound of claim 25 wherein menu from which the edit functions are selected includes a pull down menu.

27. In a system which stores digital audio on a disk drive the improvement being the use of a disk operating system which presents the structure of the disk drive as a number of virtual sectors, where each virtual sector is sized to provide a one to one correspondence between the time duration of one virtual sector and a single frame of picture.

28. An editing system for editing multiple channels of sound, such system being substantially as herein before described with reference to, and as illustrated in, the accompanying drawings.

29. A method of editing sound, such method being substantially as herein before described with reference to, and as illustrated in, the accompanying drawings.

30. A method for generating a digital audio sampling rate for use with a system having a video frequency, such method being substantially as herein before described with reference to, and as illustrated in, the accompanying drawings.

BNSDOCID: <GB___2235815A__I_> :d 1991 at The Patent Office, State House, 66/71 High Holborn, London WC1R 4TP. Further copies may be obtained from Sales Branch, Unit 6, Nine Mile Point, Cwmfelinfach, Cross Keys, Newport, NP1 7HZ. Printed by Multiplex techniques ltd, St Mary Cray, Kent